



**Kuschmann, Anja and Lowit, Anja (2018) Sentence stress in children with dysarthria and cerebral palsy. International Journal of Speech-Language Pathology. ISSN 1754-9507 ,  
<http://dx.doi.org/10.1080/17549507.2018.1444093>**

This version is available at <https://strathprints.strath.ac.uk/63354/>

**Strathprints** is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<https://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator:  
[strathprints@strath.ac.uk](mailto:strathprints@strath.ac.uk)

The Strathprints institutional repository (<https://strathprints.strath.ac.uk>) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde's intellectual output.

**Sentence stress in children with dysarthria and cerebral palsy**

**Abstract**

*Purpose:* This study aimed to advance our understanding of how children with dysarthria and cerebral palsy (CP) realise sentence stress acoustically, and how well listeners could identify the position of the stressed word within these utterances.

*Method:* Seven children with CP and eight typically-developing children participated in the experiment. Stress on target words in two sentence positions was elicited through a picture based question-answer paradigm. Acoustic parameters of stress (duration, intensity and fundamental frequency (F0)) were measured and compared between stressed and unstressed target words. For the perception experiment, ten listeners were asked to determine the position of the stressed word in the children’s productions.

*Result:* Acoustic measures showed that at group level the typically-developing children used all three acoustic parameters to mark sentence stress, whereas the children with CP showed changes in duration only. Individual performance variations were evident in both groups. Perceptually, listeners were significantly better at identifying the stressed words in the utterances produced by the typically-developing children than those of the children with CP.

*Conclusion:* The results suggest that children with CP can manipulate temporal speech properties to mark stress. This ability to modulate acoustic-prosodic features could be harnessed in intervention to enhance children’s functional communication.

**Key words:** sentence stress, cerebral palsy, dysarthria, acoustics, listener perception

## Sentence stress in children with dysarthria and cerebral palsy

In spoken English, the highlighting of particular words within an utterance is a crucial means to effective communication between conversational partners (e.g. Baltaxe & Guthrie, 1987; Fine, Bartolucci, Ginsberg, & Szatmari 1991). This phenomenon is often referred to as *sentence stress* although reference to prosodic stress, contrastive stress or focus can also be found in the literature, all essentially covering the same linguistic concept. Irrespective of the label, making some words more prominent than others enables the speaker to direct the listener to the important part of an utterance, that way structuring information in discourse and supporting listener comprehension. On the other hand, inappropriate use of stress, such as highlighting less important information or information already known to the listener can hinder the efficient processing of incoming speech and reduce comprehensibility (e.g. Terken & Neteboom, 1987).

In the speech stream, sentence stress is signalled by various acoustic-phonetic parameters, in particular increases in duration, intensity and fundamental frequency (F0) on the highlighted word (e.g. Bolinger, 1961; Lehiste, 1970, Lieberman, 1960). These acoustic parameters in turn are closely associated with the perceptual cues of length, loudness and pitch. Studies examining the relative importance of these cues identified pitch change to be the most important factor in stress perception followed by length and loudness (e.g. Fry, 1955, 1958; Morton & Jassem, 1965). In spoken English, F0 and intensity decline over the course of an utterance - presumably due to falling subglottal pressure (e.g. Cooper & Sorensen, 1981; Pierrehumbert, 1979). As a result, the acoustic realisation of a stressed word is influenced by its position within the utterance (Cooper, Eady, & Mueller, 1985; Eady & Cooper, 1986). Specifically, utterance-initial stress is signalled through a sharp fall in F0 and a compressed F0 range following the stressed word, whereas utterance-final stress results in a more gradual decline of F0 over the course of the utterance (Cooper et al., 1985). Durational and intensity

patterns are also influenced by the location of the stressed word with larger differences in both measures being observed between stressed words and their unstressed counterparts in utterance-initial position compared to stressed words positioned towards the end of utterances (Cooper et al., 1985; Huss, 1978).

Research on the marking of stress in typically-developing children suggests that children speaking West-Germanic languages including English and Dutch can generally use acoustic features to mark stress around the age of four (e.g. Hornby & Hass, 1970; Kehoe, Stoel-Gammon, & Buder, 1995; Pollock, Brammer, & Hagerman, 1993; Wonnacott & Watson, 2008). Similar to adults, children mark stressed syllables through higher F0 levels, greater intensity and longer durations (Kehoe et al., 1995; Pollock et al., 1993). However, they appear to be less consistent in terms of to what degree and in what combination they use these parameters. In addition, there is disagreement across studies at what age control over a particular acoustic feature is mastered. Kehoe and colleagues (1995) found that children as young as 18 months use duration, F0 and intensity to mark stress differences, whereas in Pollock et al.'s (1993) study children of a similar age relied on duration only. Yet another pattern was observed by Astruc, Payne, Post, del Mar Vanrell and Prieto (2012) with 2-year old children controlling F0 height but not syllable duration of words to mark stress. Wonnacott and Watson (2008) made a similar observation in 3- to 4-year old children who used F0 and intensity, but not duration to distinguish stressed from unstressed words. Based on these findings it can be concluded that children can control the same acoustic parameters as adults to mark stress, however, there is less consistency in how these are applied.

There is also evidence that the combination of parameters used by a child to mark word stress can affect how well listeners can identify the location of the stressed word. Studies by Kehoe et al. (1995) and Pollock et al. (1993) demonstrated that in instances where children changed a combination of duration, intensity and F0 listeners successfully identified

the stressed word, whereas they struggled in cases where only duration was changed to mark stress. These findings show that listeners have more difficulties locating the position of the stressed word when only a subset of acoustic parameters are manipulated.

Although most children master stress marking by the age of four, not all may be successful in this task. One group that may experience difficulties with stress marking are children with cerebral palsy (CP). CP is the most common cause of motor problems in children (Cans, De-la-Cruz, & Mermet, 2008), affecting two to three children per 1000 live births (Surman et al., 2006). It is a non-progressive disorder of movement and posture, which is caused by injury to the developing fetal or infant brain (Rosenbaum, Paneth, Leviton, Goldstein, & Bax, 2006). The motoric deficits are frequently accompanied by cognitive and sensory impairment as well as communication difficulties, whereby dysarthria is considered the most frequent communication impairment (Parkes, Hill, Platt, & Donnelly, 2010). Speech characteristics include shallow, irregular breathing, harsh and/or breathy voice with inadequate volume levels, reduced pitch variation, hypernasality, slow speech rate and imprecise articulation (Ansel & Kent, 1992; Hodge & Wellman, 1999; Nip, 2012; Nordberg, Miniscalco, & Lohmander, 2014; Workinger & Kent, 1991), affecting at least one – but often all – speech subsystems, i.e. breathing, phonation, resonance as well as articulation.

To date, no studies have reported on the acoustic features of stress production in children with CP and dysarthria. Studies on non-speech tasks examining vocal control showed a narrowed F0 range and reduced sound durations (Wit, Maassen, Gabreels, & Thoonen, 1993). Listener perception studies further showed that caregivers were able to differentiate between a number of vowel productions of children with severe dysarthria and CP that were signalled through duration and F0 changes (Patel & Salata, 2006; Whitehill, Patel, & Lai, 2008), indicating that the children seemed to be able to use these parameters to some extent to signal meaning at vowel level. Patel, Hustad, Connaghan and Furr (2012) investigated the

relation between the use of acoustic parameters and intelligibility in two- to seven-word utterances produced by children with CP. They observed that the children produced slower utterances with more variable F0 contours than children without dysarthria. Intelligibility of words improved when they were produced with heightened F0, but decreased when F0 and duration exceeded the range within which children without dysarthria produced the same words. The study provides valuable information on how children with CP and dysarthria produce acoustic parameters at sentence level. However, it remains unclear whether and to what extent children with dysarthria due to CP can purposefully manipulate these parameters to express linguistic function as would be required for the marking of sentence stress.

The only studies available on the marking of sentence stress in CP and dysarthria examined adult speech (Patel, 2004; Patel & Campellone, 2009). They showed that speakers could signal stress within short utterances by manipulating all three acoustic parameters, however, they relied more heavily on duration compared to the control speakers. In addition, F0 and intensity variation were reduced, but the narrowed range within which parameters were manipulated did not seem to affect the successful signalling of stress position as listeners were highly accurate in identifying target locations.

Whilst these results suggest that adults with CP can successfully modulate acoustic parameters to signal stress, we currently do not know whether children with restricted speech motor control due to CP have the same abilities. That is, it is currently unclear whether children with CP and dysarthria are able to use the acoustic parameters of duration, intensity and F0 to mark sentence stress - and thus - are able to convey linguistic function to listeners through means of prosody. The present study aims to fill this research gap.

**Purpose of the study**

The purpose of this study was twofold: The first was to increase our understanding of how children with dysarthria associated with CP acoustically realise sentence stress. The

second was to establish whether acoustic performances align with perceptual evaluations of children's sentence stress productions. Specifically, the study was designed to answer the following research questions:

- 1.) Do children with dysarthria and CP use duration, intensity and F0 to mark sentence stress, and how does their performance compare to that of typically-developing children?
- 2.) Can listeners successfully identify the stressed words in the productions of the children with dysarthria, and how does identification accuracy compare between the productions of children with dysarthria and typically-developing children?

### Method

The speech production and perception experiments were conducted as part of a study investigating a range of prosodic abilities in children with CP and dysarthria. Ethical approval to conduct the study was granted by the University's Ethics Committee. Informed consent for testing and recording of speech was given either by the parents or the children themselves, if they were older than 12.

### Participants

Eight children with dysarthria due to CP and eight typically developing children (CON) participated in the overall production and perception experiments. However, one child with CP (CP8) had to be excluded from the current data analysis as they were unable to complete the stress task. We decided to retain the data of all eight control speakers for the analysis to strengthen the representativeness of the control norms. Any subsequent reference to CP and CON groups thus refers to seven participants with CP (CP1-7) and eight control speakers (CON1-8) (see Table I). Participants were age-, gender- and dialect-matched (CP: five boys, range: 7-18 years, mean = 12 years; CON: six boys, range: 7-20, M = 11.8 years), with all participants being native speakers of the British English variety spoken in the West of Scotland. Hearing and vision was normal or adjusted-to-normal, allowing participants to

follow instructions and comply with tasks and tests in an appropriate manner. They further demonstrated adequate cognitive skills to complete the experiment along with a range of speech and language related tasks. The typically-developing children had no history of speech and language difficulties and no known developmental disorders. Three children had been diagnosed with dyskinetic CP, two with spastic CP, and two with ataxic CP. Gross Motor Function Classification System (GMFCS, Palisano et al., 1997) levels ranged from I to IV (median III). All participants with CP were diagnosed with dysarthria by their speech-language pathologist and had a history of regular speech and language therapy. Severity of dysarthria was determined by establishing the level of speech intelligibility at single word level using the Children’s Speech Intelligibility Measure (CSIM, Wilcox & Morris, 1999), which is considered the standard intelligibility measure in the UK. The CSIM items were rated by two unfamiliar listeners. Scores above 80% were judged as mild, scores between 50-80% as moderate and scores below 50% were considered to reflect severe dysarthria (Patel & Campellone, 2009). Based on this, three speakers were judged as having mild dysarthria; three speakers were judged as having moderate dysarthria; and one speaker was judged as having severe dysarthria. The mean percentage intelligibility for the children with CP was 67.7% and for the typically-developing group was 93.5%. In addition, the Verbal Motor Production Assessment for Children (VMPAC, Hayden & Square, 1999) was administered to obtain a better understanding of the children’s motor and speech characteristics. Scores for each of the measures are given in Table I.

---Insert Table I about here---

**Speech production experiment**

A picture based question-answer paradigm was employed to investigate the children’s ability to stress specific words in an utterance. This paradigm is frequently applied to investigate stress as it controls the information structure of the elicited responses (Roberts,



1996). For the purpose of the study, a set of short nominal phrases (NPs) was developed consisting of a pre-nominal adjective (colour – yellow, orange) and a noun (animal – cow, dog, goat, goose, horse) e.g. yellow cow. Short target structures were used to take into account potential respiratory control issues in the participants with dysarthria and CP. The utterances were controlled for number of syllables (in line with British English pronunciation) and syllable complexity with particular attention on avoiding consonant clusters. Both colours were combined with each noun resulting in ten NPs. Stress was elicited in two conditions: 1) on the adjective, and 2) on the noun, resulting in 20 phrases per participant. Overall, 300 phrases and 600 target words were elicited for the purposes of the study across the two speaker groups.

### **Task presentation and recording procedure**

Pictures were designed for each of the NP, and presented in a Microsoft Office PowerPoint® presentation to elicit the two different stress conditions. Both conditions were elicited separately. In a first step the participants were shown the pictures with contextual scenarios that prompted them to stress the colour (adjective) in the NP. This was followed by pictures and scenarios that elicited stress on the animal (noun). When describing the pictures, the participants were asked to embed the phrases in a carrier sentence (e.g. “The yellow cow jumped out.”). Appendix A details how the task was introduced to each child, and exemplifies how target utterances were elicited.

Audio recordings were made in each participant’s home using a portable Edirol R-09HR MP3 recorder. Sound files were saved as .wav files at a sampling rate of 44.1 kHz. Participants were seated comfortably in front of a laptop computer and asked to work through the presentation at their own pace with pictures being shown one at a time. The presentation started with a familiarisation phase including five practice stimuli. This was followed by the experimental phase. Target stimuli were interspersed with 24 filler items, i.e. the same

animals but with a different colour. Recording settings, instructions and experimental design were the same for each participant.

**Data selection, annotation and analyses**

Data were analysed acoustically using Praat speech analysis software (version 5.3.39; Boersma & Weenink, 2018) to obtain measures on duration, intensity and F0 for each NP. In a first step, the participants' responses were evaluated to exclude answers containing non-target words (e.g. duck instead of goose) and other changes (e.g. replacing the noun with a pronoun) to ensure that subsequent measures of acoustic parameters were conducted on the same words. As a result of this data cleaning process the final data set consisted of 574 target words with a mean response inclusion of 97% for the eight typically developing children (range: 85-100%) and 94% for the seven children with CP (range: 88-100%).

The selected recordings were segmented into words based on waveform and formant changes in the spectrogram (Turk, Nakai, & Sugahara, 2006), and annotated using waveform and pitch track information. For each response, four durational landmarks were labelled as well as two F0-related landmarks (illustrated in Figure 1): 1) Beginning of target word 1 and 2 (marking the onset of the word-initial phoneme for adjective and noun, respectively); 2) End of target word 1 and 2 (marking the offset of the word-final phoneme for adjective and noun, respectively); 3) 1max - the highest F0 value on target word 1 (adjective), and 4) 2max - the highest F0 value on target word 2 (noun).

With regard to the F0 landmarks, each file was inspected to identify and correct potential octave errors. Two Praat scripts were then employed to automatically extract values for the acoustic parameters *duration*, *intensity* and *F0*. For duration, the length of each target word was extracted (in ms) using the labels on tier 2; for intensity, the maximum decibel (dB) value on each labelled target word was extracted (based on durational landmarks); and for F0,

Hz maxima were extracted on each target words from the labelled points on tier 3 (see Figure 1).

---Insert Figure 1 about here---

## **Listener perception experiment**

### **Listener participants**

Ten monolingual speakers of British English (20-37 years,  $M = 23.4$ , 8 female) served as listeners for the perception experiment. One listener was a speech-language pathologist (SLP) graduate, and nine were SLP students in their 3<sup>rd</sup> or 4<sup>th</sup> year. Whilst all were experienced with dysarthric speech, they were not familiar with the presented stimulus materials.

### **Perception experiment set-up**

The perception experiment was conducted in a quiet room on University premises. Listeners were seated in front of a laptop computer, with E-Prime 2.0 Professional software (SP1 2.0.10.353; Psychology Software Tools, Pittsburgh, PA) being used to present the stimuli. The software interface guided the listeners through instructions, practice and target stimuli, and recorded their responses. Stimuli were presented through Sennheiser HD 202 headphones with a pre-set, comfortable volume level. Each listener heard 248 utterances<sup>1</sup> by the typically-developing children and the children with CP, which were randomly presented in blocks of 50. Each block was followed by a break, with additional breaks possible at any time during the experiment. Target stimuli were interspersed with 24 filler items, resulting in 272 stimuli presented. Each listener judged the whole data set. A forced choice paradigm was employed to determine the position of sentence stress, i.e. for each utterance listeners were asked to decide whether the speaker stressed the adjective (i.e. colour), or the noun (i.e.

---

<sup>1</sup> This number is lower than the set used for the acoustic analysis as target structures that were unintelligible unfamiliar speakers were not included in the perception experiment.

animal). Listeners could repeat each stimulus twice before making their decision. All listeners completed the experiment in about 45 to 60 minutes.

Intra-rater agreement was determined for each listener by randomly repeating 28 recordings, i.e. 10% of the data. A listener was considered reliable, if they responded consistently to original and repeated stimulus, irrespective of whether the response was correct or not. Mean intra-rater agreement for the listeners was 82.7%; that is on average 25 of the 28 stimuli were reliably identified (range: 76.7% - 96.67%).

**Statistical analyses**

A number of statistical tests were employed to answer the research questions. The level of significance was set at  $<0.05$ .

To examine the use of the acoustic parameters to mark stress (research question 1), a series of mixed model ANOVAs were conducted for each parameter, i.e. duration, intensity and F0 (dependent measures). *Stress condition* (two levels: stressed, unstressed) and *stress position* (two levels: word 1 (adjective), word 2 (noun)) served as within-subject factors, and group (CP and CON) as between-subject factor. In addition, two types of pairwise comparisons within each group were conducted using the non-parametric Wilcoxon signed-rank test to determine whether each group differentiated acoustically between stressed and unstressed words. First, for each of the specified dependent measures i.e. duration, intensity and F0, comparisons *across* utterance pairs were conducted (i.e. stressed vs. unstressed version of adjective and noun). Second, comparisons *within* utterance pairs (i.e. from stressed adjective to unstressed noun, and vice versa) were carried out for intensity and F0 to establish acoustic changes from one target word to the other. Due to the inherent differences in length between the adjective and noun, duration was not included in this comparison.

Finally, the data was normalised to evaluate individual performance differences with regard to the realisation of stressed and unstressed words across utterances. For each speaker

the mean duration, intensity and F0 of stressed and unstressed adjectives and nouns were calculated. In a next step, it was established to what extent each word differed from the mean, with expectations that the values for the stressed words would be above the mean and for unstressed words below the mean.

A mixed ANOVA was conducted to answer research question 2, which examined listeners' abilities to identify the position of the stressed word. *Stress position* (adjective vs. noun) served as within-subject factor, and group (CP and CON) as between-subject factor.

### **Intra- and inter-rater reliability**

Intra-rater performance for the acoustic measures of duration, intensity and F0 were established on 10% of the data using Pearson correlation coefficient (two-tailed). The results revealed significant correlations between measures (duration  $r = .935$ , intensity  $r = 0.991$ , F0  $r = 0.990$ ). Inter-rater reliability was measured using the mean intraclass correlation coefficient (ICC), and showed excellent agreement between raters across all three acoustic measures (duration: 0.97, 95% CI 0.94-0.98; intensity: 0.995, 95% CI 0.992-0.997; F0: 0.995, 95% CI 0.991-0.997). This was reflected in the small differences in mean and standard deviations of the absolute values for duration (22.52ms (23.23ms)), intensity (0.06dB (0.03dB)) and F0 (0.82Hz (0.68Hz)) in the inter-rater analysis. Similarly negligible differences were found for the intra-rater analysis (duration: 0.55ms (3.94ms); intensity: 0.17dB (0.12dB) and F0: 0.62Hz (0.77Hz).

### **Result**

Table II summarises the results of the mixed model ANOVA for the within-subject effects *stress condition* (stressed vs. unstressed) and *stress position* (adjective vs. noun), the between-subject effect *group* and the interaction between within- and between-subject effects for each parameter.

---Insert Table II about here---

**Duration**

Table II shows that there was a significant main effect of group for duration. A significant main effect was also obtained for stress condition. The interaction effect (condition\*group) was not significant. Wilcoxon signed rank tests across utterance pairs showed significant differences between stressed and unstressed words for both groups in both positions (CON, word 1:  $Z = -3.53, p = 0.000$ ; word 2:  $Z = -4.46, p = 0.000$ ; CP: word 1:  $Z = -2.94, p = 0.003$ ; word 2:  $Z = -2.32, p = 0.020$ ). In sum, although the children with CP generally spoke more slowly, both groups used word duration to differentiate between stressed and unstressed words, with stressed words being significantly longer than their unstressed counterparts in both stress positions (see Figure 2).

---Insert Figure 2 about here---

**Intensity**

In contrast to duration there was no significant main effect of group for intensity (see Table II). Tests of within-subject effects showed significant main effects for stress condition and position. A significant interaction between variables was observed for position and group but not for condition and group. Wilcoxon signed rank tests *across* utterance pairs revealed significant differences between stressed and unstressed words for the noun position for the CON group, as well as a trend in this direction for the adjective position (CON, adjective:  $Z = -1.90, p = 0.058$ ; noun:  $Z = -2.71, p = 0.007$ ). For the CP group no significant difference between the productions of stressed and unstressed counterparts was observed (adjective:  $Z = -0.26, p = 0.793$ ; noun:  $Z = -1.12, p = 0.523$ ). These results suggest that the CON group, but not the CP group, manipulated intensity levels to some extent to mark stress, with stressed nouns being significantly louder than their unstressed counterparts (see Figure 3).

---Insert Figure 3 about here---

Additionally, comparisons *within* utterance pairs were conducted to evaluate the change of intensity levels from adjective to noun position in both groups. Results showed a significant intensity drop in the stressed-unstressed condition for the CON group ( $Z = -4.75$ ,  $p = 0.000$ ; dB drop = 2.99dB), but no significant rise in the unstressed-stressed scenario ( $Z = -0.338$ ,  $p = 0.735$ , dB rise = 0.66dB). For the CP group, a significant intensity decrease was observed in both conditions (stressed to unstressed:  $Z = -4.71$ ,  $p = 0.000$ , dB drop = 3.65dB; unstressed to stressed:  $Z = -3.78$ ,  $p = 0.000$ , dB drop = 2.88dB). These results suggest that the children with CP consistently produced the adjective in each utterance pair more loudly irrespective of whether it was stressed or unstressed, whereas the CON group showed this pattern only in the expected stressed adjective to unstressed noun condition.

## F0

The analysis of F0 resulted in a significant main effect of group as well as stress condition, but not for position (see Table II). In addition, a significant interaction was observed between stress condition and group but not position and group. Wilcoxon signed rank tests *across* utterance pairs showed significant differences between stressed and unstressed words for the CON group (CON, word 1:  $Z = -3.53$ ,  $p = 0.000$ ; word 2:  $Z = -3.84$ ,  $p = 0.000$ ), but not for the CP group (word 1:  $Z = -0.99$ ,  $p = 0.320$ ; word 2:  $Z = -1.85$ ,  $p = 0.065$ ). This means that the control group, but not the children with CP, produced stressed words that were significantly higher in F0 than those in unstressed position. The difference for the CON group was more pronounced on the adjective (see Figure 4).

---Insert Figure 4 about here---

Subsequent comparisons *within* utterance pairs showed a significant drop in F0 in stressed-unstressed sequences for the CON group ( $Z = -3.52$ ,  $p = 0.000$ ; F0 drop = 21.02Hz) as well as a significant increase from unstressed to stressed word pairs ( $Z = -3.64$ ,  $p = 0.000$ ; F0 rise = 23.7Hz). Thus, the CON group manipulated F0 levels within word pairs to mark the

position of the stressed word. This could not be observed for the CP group, which produced a significant F0 fall on the stressed to unstressed pair ( $Z = -3.54, p = 0.000$ ; F0 drop = 10.37Hz), but no significant change for the unstressed to stressed pair ( $Z = -.012, p = 0.991$ ; F0 drop = 0.64Hz).

Overall, the control speakers employed duration and F0 systematically to mark stress position with stressed words being significantly longer and higher in F0 than their unstressed counterparts. They also employed intensity to some extent to signal stress position within and across utterances. In contrast, the children with CP only manipulated duration in a comparable way.

**Individual performances**

In light of the variation in type of CP and severity of dysarthria individual performances were examined to better understand behaviour trends within each group. Table III shows for each group and individual the extent to which the stressed words differed from the normalised mean. Most children in both groups employed duration to mark stress, with stressed words being on average 5.3% (CON) and 4.9% (CP), respectively, longer than the mean, reflecting findings from the group analyses. The individual results confirm that intensity was not used consistently to mark sentence stress, with two typically-developing children (CON5, CON7) and five children with CP (CP2, CP3, CP5, CP6, CP7) yielding negative results, indicating that their stressed words were quieter than the mean. Only CON1 and CP1 appeared to use intensity to differentiate between stressed and unstressed words in a systematic way. In terms of F0, findings are mixed with two typically-developing children (CON1, CON8) and two children with CP (CP1, CP4) showing percentage increases of more than 4% on the stressed word whereas for others a negative value indicated that the stressed words were lower in F0 than the unstressed words (CON7 vs. CP2, CP3, CP6).

---Insert Table III about here---



### Listener accuracy by speaker group and stress position

For the control group listeners were 72.71% (SD: 5.02%) accurate at identifying the position of the stressed word across all utterances compared to 60.66% (SD: 4.87%) for the children with CP. The magnitude of difference between groups was 12.04%; with listener accuracy ranging from 66 to 82% for the typically-developing children and from 55 to 68% for the children with CP. The difference in the listeners' abilities to identify the position of the stressed word for both groups was statistically significant ( $F(1, 37.624)$ ,  $p < 0.001$ ), suggesting that listeners were more accurate at determining the position of the stressed word in the utterances of the typically-developing children.

In terms of stress position, listener accuracy for the productions of the typically-developing children was 72.29% for word 1 (colour) and 73.14% for word 2 (animal), and 62.08% and 59.25% for the children with CP respectively. Across all productions together, listeners accurately identified 67.18% (SD: 5.52%) of stressed words in position 1 (colour), and 66.19% (SD: 5.89%) in position 2 (animal). The difference was not statistically significant ( $F(1, 0.137)$ ,  $p = 0.720$ ), indicating that stress position did not have a role in listeners' accuracy scores.

### Discussion

The current study investigated how children with dysarthria due to CP and typically-developing children use the acoustic parameters of duration, intensity and F0 to signal sentence stress. We also aimed to understand the link between the use of these parameters and listeners' abilities to identify the target word within these utterances. Findings indicated that the group of children with CP manipulated duration to mark sentence stress, but did not employ intensity and F0. In contrast, the typically-developing children employed a combination of all three parameters to signal sentence stress. The perception experiment revealed that listeners were significantly better at determining the stressed word in the

productions of the typically-developing children than in those of the children with CP. Stress position did not have an effect on perception accuracy in either group.

**Use of acoustic parameters to mark sentence stress**

The durational results suggest that the children with CP, like the typically-developing children, employed word duration to differentiate between stressed and unstressed words. Their successful use of duration to mark stress in short utterances is in line with findings from previous studies on the use of durational cues in vocalisations by children with dysarthria and CP (Patel & Salata, 2006; Whitehill et al., 2008). The current findings further reflect observations made by Patel and Campellone (2009), who found that adult speakers with CP not only used duration, but relied most heavily on this parameter to mark contrastive stress in short utterances. On the other hand, our findings show that the children with CP did not reliably employ F0 and intensity to mark sentence stress, which is in contrast to the observations made for adult speakers, who manipulated F0 and intensity albeit within a reduced range (Patel & Campellone, 2009). Taking these observations into account, our finding may either suggest that children with CP may not be able to manipulate F0 and intensity on stretches of speech that go beyond segmental level, or that they have yet to reach adult-like performance for these parameters. At the same time, it is important to highlight that the typically-developing children recorded in the present study did not employ intensity in a consistent manner either, suggesting that intensity was the least reliable parameter employed by both groups. A similar observation was made by Patel et al. (2012) who found that differences in duration and F0 realisations had an impact on speech intelligibility, whereas intensity features did not separate high and low intelligibility words in either group investigated. These and our observations tie in with the fact that loudness as a perceptual cue is deemed less relevant for stress marking than pitch and length, and it may be argued that they reflect children's awareness of the relative importance of the different perceptual cues.

Research on acquired motor speech disorders e.g. ataxic dysarthria have reported similar observations with intensity manipulations in stress tasks (Lowit, Kuschmann, MacLeod, Schaeffler, & Mennen, 2010). Therefore, it may be possible that this parameter might be particularly susceptible to problems in motor control irrespective of whether they are acquired or developmental in nature. On the other hand, the lack of intensity modulation observed in the typically-developing children could also point to the fact the current results were to some degree influenced by the maturity of the speech mechanism. However, the counterargument to this assumption is the fact that the one control speaker who did show noticeable changes in intensity was one of the youngest participants. The current results thus reflect some of the inconsistencies reported in previous research on typically-developing children in relation to which parameters were being used to signal stress. For the moment, the group as well as individual results indicate that duration appeared to be the most frequently and extensively applied acoustic parameter to mark stress, followed by F0 and intensity.

### **Perceptual evaluation of stress marking**

The perception experiment aimed at determining whether listeners could identify the stressed word within the short utterances produced by the typically-developing children and children with CP. The results showed that the listeners were less successful in determining stress position in the productions by the children with CP, who made less consistent use of the range of acoustic parameters available to mark sentence stress. This finding ties in with evidence from Kehoe et al. (1995) and Pollock et al. (1993) that insufficient use or limited control over acoustic parameters can lead to difficulties in locating stress position. Given that the listening task was based on a binary choice paradigm which suggests a 50% chance of guessing the correct target, the 61% success rate is indicative of the listeners' struggle to identify the position of the stressed word. However, this not only seems to hold true for the children with CP, but the typically-developing children as well. Whilst the difference between

groups in terms of listener accuracy was significant, in about one in every four productions by the typically-developing children stress position was not identified correctly. This might in part be due to the fact that the control group did not consistently use a combination of all cues either, as reflected in the group results for intensity as well as individual performances. Speaker CON7, for instance, only employed duration (see Table III), and feedback from listeners on this particular speaker showed that they struggled to determine stress location as she spoke quite fast, possibly limiting the usefulness of duration as a cue, and in addition not making use of intensity and F0 to compensate for this.

The perceptual findings also revealed that the position of the stressed word did not influence perception accuracy. That is, neither the control group nor the children with CP appeared to have produced stress on colour or animal differently, rendering both sentence positions equally amenable to prosodic manipulations. Having said that the children with CP showed a trend towards higher F0 and intensity on the first word, i.e. the adjective, irrespective of its stress status. This could signal that they experienced motoric restrictions and could only follow the typical pattern of declination, i.e. higher F0/intensity at the start with a gradual lowering across utterances, which was appropriate for the stressed-unstressed scenario, but not the unstressed-stressed condition which requires more control over these features. Raising F0 entails purposeful physiological manipulation to increase vocal fold tension through changes in the level of activation of the intrinsic laryngeal muscles and/or a rise in subglottal pressure through increase in glottal airflow (Ladefoged, 1963). Given the frequent issues with phonatory/respiratory control in people with dysarthria this physiological mechanism may be less efficient or may operate differently, rendering sentence position a potentially robust factor for predicting prosodic performance in speakers with dysarthria (Cooper et al., 1985; Eady & Cooper, 1986; Huss, 1978). On the other hand, this did not

appear to affect perception accuracy, which was similar for each position, suggesting again that the use of duration may have compensated for this to at least some extent.

### **Clinical implications**

Children with CP do make use of the temporal parameter to mark sentence stress and this strategy appears to be effective to some extent. At the same time, our findings showed that the typically-developing children, who used a greater variety of parameters, were more successful in getting their message across. Acknowledging the observed individual variation in the use of cues, the results of our study suggest that there may be some ability to manipulate acoustic parameters to signal sentence stress in almost every child with dysarthria and CP. Identifying the functional cues and building on these may therefore be considered a viable treatment strategy to enhance communicative function in this group. In terms of remediation strategies two approaches appear promising. Firstly, treatment could focus on refining the ability to manipulate temporal aspects of speech to a greater extent that way compensating for not using F0 and intensity to the same degree as the typically-developing children. In combination with pausing as a durational cue this could improve listeners' understanding of children's utterances in spoken interactions. The second option would be to support and encourage the use of the parameters F0 and intensity, whereby priority should be given to the modelling of F0 contours. This may allow the children to approach adult-like performance as observed by Patel and Campellone (2009), where the use of these parameters was functional within a limited range and as a result yielded a ceiling effect in terms of accurate perception of stress location. Overall, our findings do make a case for including prosodic exercises more readily in clinical practice as one option for children with CP to improve communicative function. Intervention studies are needed to assess the potential of these strategies.

### **Conclusion**

The current study has made novel and valuable contributions towards understanding functional prosodic abilities in a small group of children with dysarthria and CP at utterance level. It describes children's ability to manipulate acoustic parameters and highlights the importance of adequate control over these for the perception of stress marking. Whilst the findings can be considered robust - marked differences between both groups in terms of production and perception were observed - a detailed investigation with a larger sample of children would be beneficial to replicate and extend findings, and to account for individual differences inherent in the present sample in terms of severity and type of dysarthria as well as age range. Also, whilst scripted speech had the marked advantage to elicit the same words across contexts in a controlled environment, which was important for the current investigation, it would be beneficial to explore how children employ the different acoustic parameters in more natural contexts.

In conclusion, this study was the first to investigate sentence stress in a group of children with CP and dysarthria. We have demonstrated that these children used duration, but not intensity and F0 to mark sentence stress, compared to their typically-developing peers who employed all three parameters. As a result of the differences in the use of acoustic parameters, the children with CP were less successful in signalling the position of the stressed word to listeners, suggesting a clear link between the use of acoustic properties and listener success. Our results also highlight the fact that there may be some ability in every child to manipulate acoustic parameters to some extent. Therefore, children may benefit from intervention targeted at identifying and building on these skills. Specific prosodic intervention may help to improve children functional communication with regard to stress marking and could constitute an important stepping stone to enhancing children's overall communicative success.

## Acknowledgments

This research was supported by a British Academy postdoctoral fellowship (PF120045) awarded to the first author. We would like to thank the participants and their families for their time and enthusiasm for the study as well as Capability Scotland, Race Running Scotland and Bobath Scotland for their invaluable help with recruiting participants.

## Declaration of interest

The authors report no declarations of interest.

References

- Ansel B. M., & Kent, R. D. (1992). Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *Journal of Speech and Hearing Research*, 35, 296–308.
- Astruc, L., Payne, E., Post, B., del Mar Vanrell, M., & Prieto, P. (2012). Tonal Targets in Early Child English, Spanish, and Catalan, *Language and Speech*, 56(2), 229–253.
- Baltaxe, C., & Guthrie, D. (1987). The use of primary sentence stress by normal, aphasic, and autistic children. *Journal of Autism and Developmental Disorders*, 17, 255–271.
- Boersma, P., & Weenink, D. (1992–2018). Praat – Doing Phonetics by Computer, Version 5.3.39 [Computer software]. Retrieved from [www.praat.org](http://www.praat.org)
- Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language*, 37, 83–96.
- Cans, C., De-la-Cruz, J., & Mermet, M. (2008). Epidemiology of cerebral palsy. *Paediatrics and Child Health*, 18(9), 393–398.
- Cooper, W. E., & Sorensen J. M. (1981). *Fundamental frequency in sentence production*. Springer-Verlag: New York/Berlin.
- Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question–answer contexts. *Journal of the Acoustical Society of America*, 77, 2142–2156.
- Eady, S. J. & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, 80, 402–415.
- Fine, J., Bartolucci, G., Ginsberg, G., & Szatmari, P. (1991). The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, 32, 771–782.
- Fry, D. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27, 765–768.
- Fry, D. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 126–152.



- Hayden, D., & Square, P. (1999). *VMPAC manual*. San Antonio: The Psychological Corporation.
- Hodge, M. M., & Wellman, L. (1999). Management of children with dysarthria. In A. J. Caruso & E. Strand (Eds.), *Clinical management of motor speech disorders in children* (pp. 209-280). New York: Thieme.
- Hornby, P. A., & Hass, W. A. (1970). Use of Contrastive Stress by Preschool Children. *Journal of Speech, Language, and Hearing Research, 13*, 395-399.
- Huss, V. (1978). English word stress in the post-nuclear position. *Phonetica, 35*, 86-105.
- Kehoe, M., Stoel-Gammon, C., & Buder, E. H. (1995). Acoustic Correlates of Stress in Young Children's Speech. *Journal of Speech, Language, and Hearing Research, 38*, 338-350.
- Ladefoged, P. (1963). Some physiological parameters in speech. *Language and Speech, 6*, 109-119.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America, 32*, 451-454.
- Lowit, A., Kuschmann, A., MacLeod, J., Schaeffler, F., & Mennen, I. (2010). Sentence stress in ataxic dysarthria-A perceptual and acoustic study. *Journal of Medical Speech-Language Pathology, 18*, 77-82.
- Morton, J., & Jassem, W. (1965). Acoustic correlates of stress. *Language & Speech, 8*, 159-181.
- Nip, I. (2012). Kinematic characteristics of speaking rate in individuals with cerebral palsy: A preliminary study. *Journal of Medical Speech-Language Pathology, 20*, 88-94.
- Nordberg, A., Miniscalco, C., & Lohmander, A. (2014). Consonant production and overall

- speech characteristics in school-aged children with cerebral palsy and speech impairment. *International Journal of Speech-Language Pathology*, 16(4), 386–395.
- Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., & Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine & Child Neurology*, 39, 214–223.
- Parkes, J., Hill, N., Platt, M. J., & Donnelly, C. (2010). Oromotor dysfunction and communication impairments in children with cerebral palsy: a register study. *Developmental Medicine & Child Neurology*, 52(12) 1113 – 1119.
- Patel, R. (2004). The Acoustics of Contrastive Prosody in Adults With Cerebral Palsy. *Journal of Medical Speech-Language Pathology*, 12, 189-193.
- Patel, R., & Campellone, P. (2009). Acoustic and Perceptual Cues to Contrastive Stress in Dysarthria. *Journal of Speech, Language, and Hearing Research*, 52, 206-222.
- Patel, R., Hustad, K. C., Connaghan, K. P., & Furr, W. (2012). Relationship Between Prosody and Intelligibility in Children with Dysarthria. *Journal of Medical Speech-Language Pathology*, 20(4), 95-99.
- Patel, R., & Salata, A. (2006). Using computer games to mediate caregiver-child communication for children with severe dysarthria. *Journal of Medical Speech Language Pathology*, 14, 279–284.
- Pierrehumbert, J. B. (1979). The perception of fundamental frequency declination. *The Journal of the Acoustical Society of America*, 66, 363–369.
- Pollock, K., Brammer, D., & Hagerman, C. (1993). An acoustic analysis of young children's productions of word stress. *Journal of Phonetics*, 21, 183-203.
- Roberts, C. (1996). Information structure: Towards an integrated theory of formal pragmatics. In J. H. Yoon & A. Kathol (Eds.), *OSU Working Papers in Linguistics*, Vol. 49. Columbus: The Ohio State University. Department of Linguistics.

- Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., & Bax, M. (2006). A report: the definition and classification of cerebral palsy. *Developmental Medicine & Child Neurology*, 49, 8-14.
- Surman, G., Bonellie, S., Chalmers, J., Colver, A., Dolk, H., Hemming, K., ... Platt, M. J. (2006). UKCP: a collaborative network of cerebral palsy registers in the United Kingdom. *Journal of Public Health*, 28(2), 148-156.
- Terken, J., & Nöteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, 2, 145-163.
- Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In S. Sudhoff (Ed.), *Methods in Empirical Prosody Research* (pp. 1-28). Berlin, Germany and New York, NY: Walter De Gruyter.
- Whitehill, T., Patel, R., & Lai, J. (2008). The use of prosody by children with severe dysarthria: A Cantonese extension study. *Journal of Medical Speech Language Pathology*, 16, 293-301
- Wilcox, K., & Morris, S. (1999). *Children's Speech Intelligibility Measure*. Pearson.
- Wit, J., Maassen, B., Gabreels, F. J. M., & Thoonen, G. (1993). Maximum performance tests in children with developmental spastic dysarthria. *Journal of Speech and Hearing Research*, 36, 452-459.
- Wonnacott, E., & Watson, D. G. (2008). Acoustic emphasis in four year olds. *Cognition*, 107(3), 1093-1101.
- Workinger, M. S., & Kent, R. D. (1991). Perceptual analysis of the dysarthrias in children with athetoid and spastic cerebral palsy. In C. A. Moore, K. M. Yorkston, & D. R. Beukelman (Eds.), *Dysarthria and Apraxia of Speech: Perspectives on Management* (pp. 109-126). Baltimore, MD: Paul Brookes.

Table I

Demographic data for each participant (CP - cerebral palsy, CON – control) including GMFCS, CSIM, and VMPAC scores

Child	Sex	Age	Type of CP	GMFCS	CSIM % intelligible single words	Dysarthria severity	Global Motor	Focal Oro- motor	Sequencin g	Connected Speech/ Language	Speech Characteris tics
CP1	m	7	dyskinetic	I	80	mild	70% (sev)	79% (sev)	96% (WNL)	84% (sev)	86% (mod)
CP2	m	7	spastic	IV	91	mild	80% (sev)	88% (sev)	63% (sev)	84% (sev)	86% (mod)
CP3	m	16	spastic	III	52	moderate	70% (sev)	78% (sev)	96% (WNL)	87% (sev)	86% (mod)
CP4	m	18	ataxic	IV	69	moderate	65% (sev)	66% (sev)	96% (WNL)	80% (sev)	57% (sev)
CP5	m	13	ataxic	III	38	severe	55% (sev)	54% (sev)	80% (sev)	67% (sev)	57% (sev)
CP6	f	8	dyskinetic	III	60	moderate	50% (sev)	77% (sev)	96% (WNL)	87% (sev)	43% (sev)
CP7	f	15	dyskinetic	III	84	mild	65% (sev)	88% (sev)	100% (WNL)	93% (mild)	86% (mod)
CON1	m	7	-	-	88	-	-	-	-	-	-
CON2	m	8	-	-	83	-	-	-	-	-	-
CON3	m	16	-	-	99	-	-	-	-	-	-
CON4	m	20	-	-	100	-	-	-	-	-	-
CON5	m	14	-	-	98	-	-	-	-	-	-
CON6	f	7	-	-	94	-	-	-	-	-	-
CON7	f	16	-	-	92	-	-	-	-	-	-
CON8	m	6	-	-	94	-	-	-	-	-	-

**Note:** f - female, m - male; GMFCS - Gross Motor Function Classification System (I - walks without limitations, II - walks with limitations, III - walks using a hand-held mobility device, IV - self-mobility with limitations); CSIM - Children’s Speech Intelligibility Measure (mild - 80% and above, moderate - 50 to 80%, severe - below 50%); VMPAC - Verbal Motor Production Assessment for Children (WNL - Within Normal Limits, mild - mild, mod - moderate, sev - severe)

Table II

Mixed model ANOVA results by acoustic parameter (stress position for duration not tested due to intrinsic differences in length of adjective and noun)

	Duration			Intensity			F0		
	<i>df</i>	<i>F</i>	<i>p</i>	<i>df</i>	<i>F</i>	<i>p</i>	<i>df</i>	<i>F</i>	<i>p</i>
Mixed model ANOVA									
Between-subject effects									
Group	1	46.89	<0.001	1	130.96	0.289	1	7.35	0.008
Within-subject effects									
Stress condition	1	55.72	<0.001	1	12.27	0.001	1	22.46	<0.001
Stress position	N/A	N/A	N/A	1	47.48	<0.001	1	0.09	0.768
Interaction effects									
Condition*group	1	0.023	0.879	1	3.46	0.065	1	6.65	0.011
Position*group	N/A	N/A	N/A	1	8.86	0.004	1	0.291	0.591

Table III

Percent difference of stressed words from normalised mean for both speaker groups (CON - control, CP - cerebral palsy)

	<i>Duration</i>	<i>Intensity</i>	<i>F0</i>		<i>Duration</i>	<i>Intensity</i>	<i>F0</i>
CON1	8.43	7.08	12.98	CP1	7.01	3.54	4.89
CON2	0.21	0.01	3.09	CP2	7.97	-0.47	-3.52
CON3	7.67	0.89	0.64	CP3	2.84	-0.53	-0.23
CON4	3.95	0.14	1.84	CP4	4.76	1.34	4.08
CON5	4.40	-0.78	0.06	CP5	1.20	-0.18	1.51
CON6	4.96	1.74	2.35	CP6	10.42	-1.07	-0.41
CON7	4.80	-0.21	-0.56	CP7	2.09	-0.53	0.33
CON8	7.95	1.23	5.81				
Mean	5.30	1.26	3.28	Mean	4.88	0.30	0.95
SD	2.71	2.49	4.40	SD	3.63	1.61	2.86

## Figure captions

Figure 1: The utterance “Yellow cow’s jumped out” - produced by speaker CON1 with stress on colour - illustrates the different annotation levels including durational and F0 landmarks. The Figure further shows the oscillogram (sound wave), the spectrogram (frequency distribution) and the pitch contour (in blue).

Figure 2: Mean duration of stressed and unstressed target words in the different stress positions (adjective and noun) for both speaker groups (CON - control, CP - cerebral palsy)

Figure 3: Mean intensity of stressed and unstressed target words in the different stress positions (adjective and noun) for both speaker groups (CON - control, CP - cerebral palsy)

Figure 4: Mean F0 of stressed and unstressed target words in the different stress positions (adjective and noun) for both speaker groups (CON - control, CP - cerebral palsy)

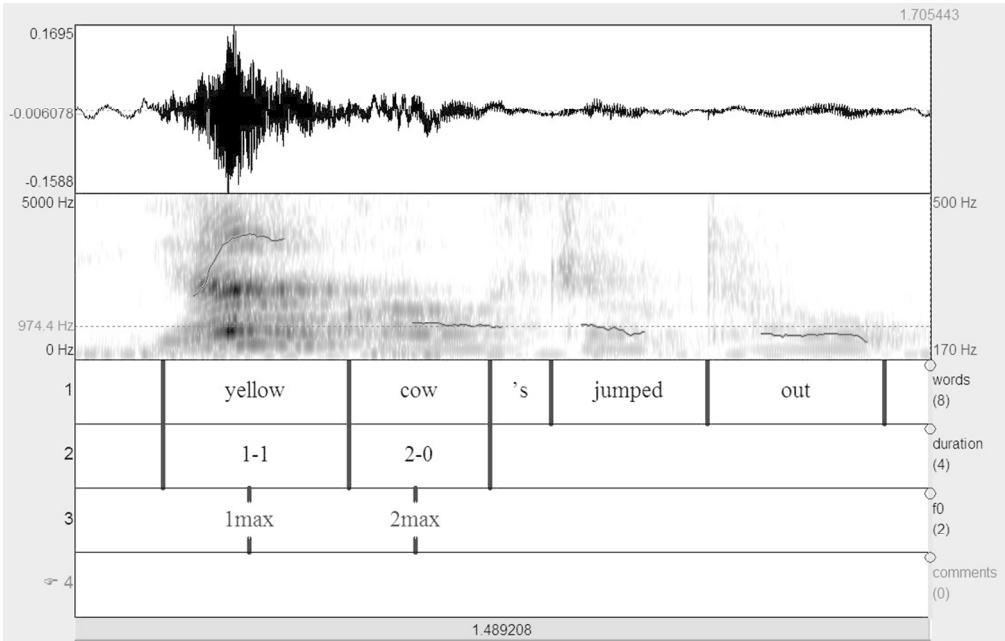


Figure 1

94x60mm (300 x 300 DPI)



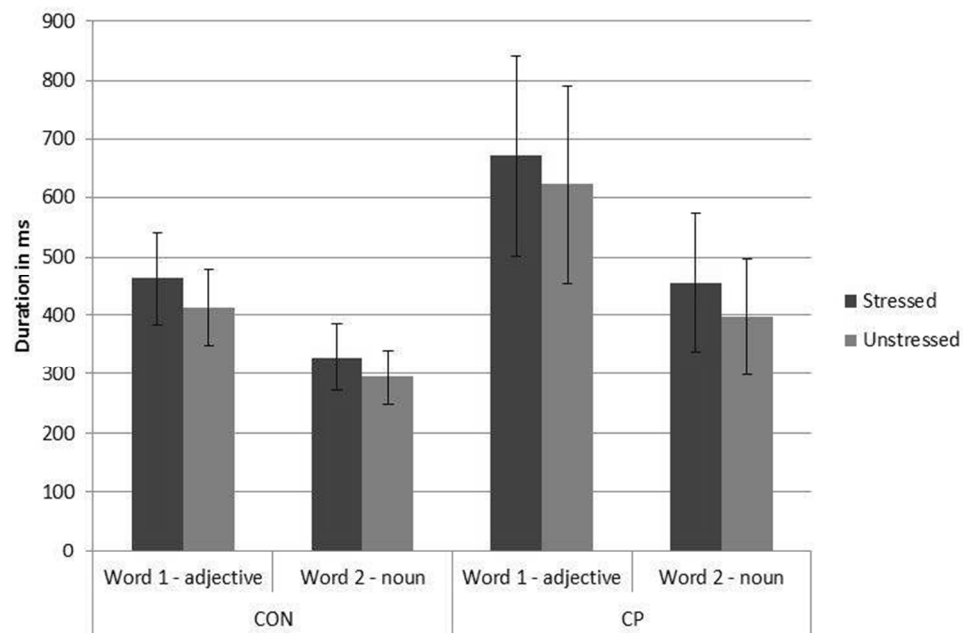


Figure 2

65x42mm (300 x 300 DPI)

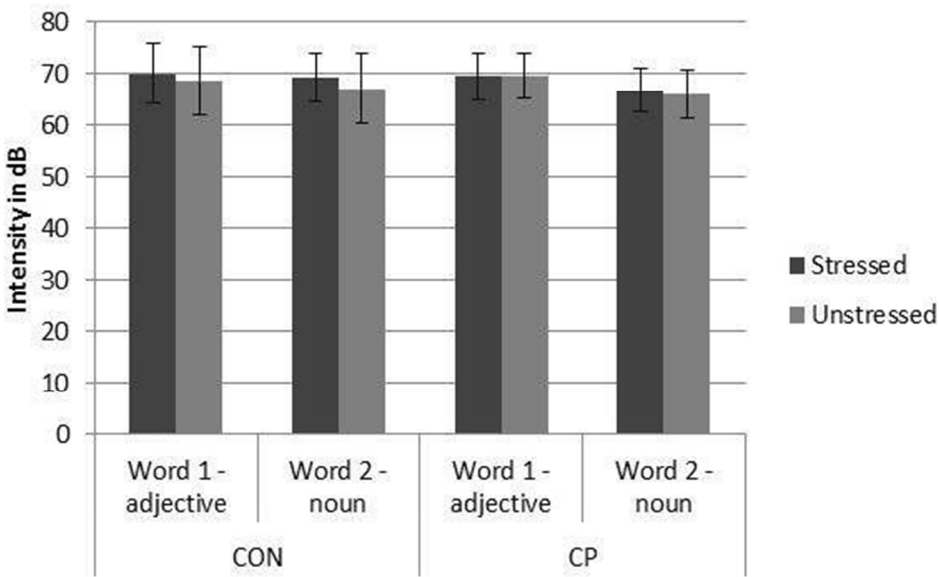


Figure 3  
50x30mm (300 x 300 DPI)

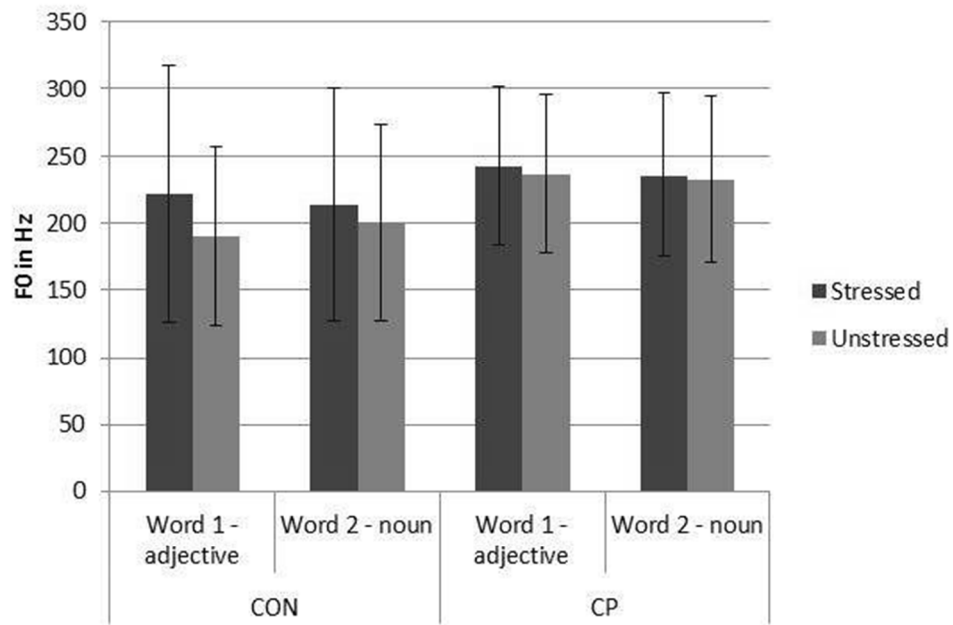


Figure 4

54x35mm (300 x 300 DPI)

**Focus task - Script**

**“I am going to show you some pictures with farm animals. They are quite funny as they have all different colours”**

[Open ppt and let the child look at the first picture for a few moments – test 1]

**“Here is a green dog, a grey dog and a yellow dog.”**

[Point to the picture]

**“They are a bit bored and keep jumping out. They take turns jumping over the fence.”**

[next slide – test 2]

**“Here, the green dog jumped out.”**

[next slide – test 4]

**“And here...which dog jumped out?”**

[wait for answer – correct if necessary, say that should answer in a sentence]